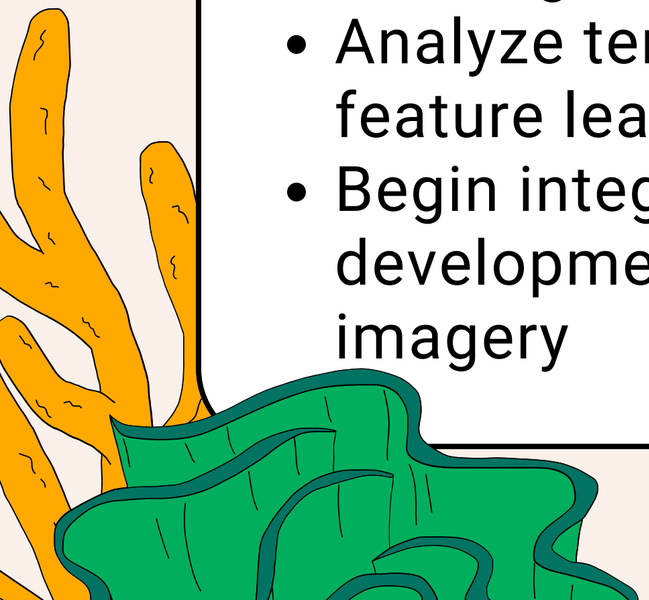
Harmful Algae Bloom Detection System

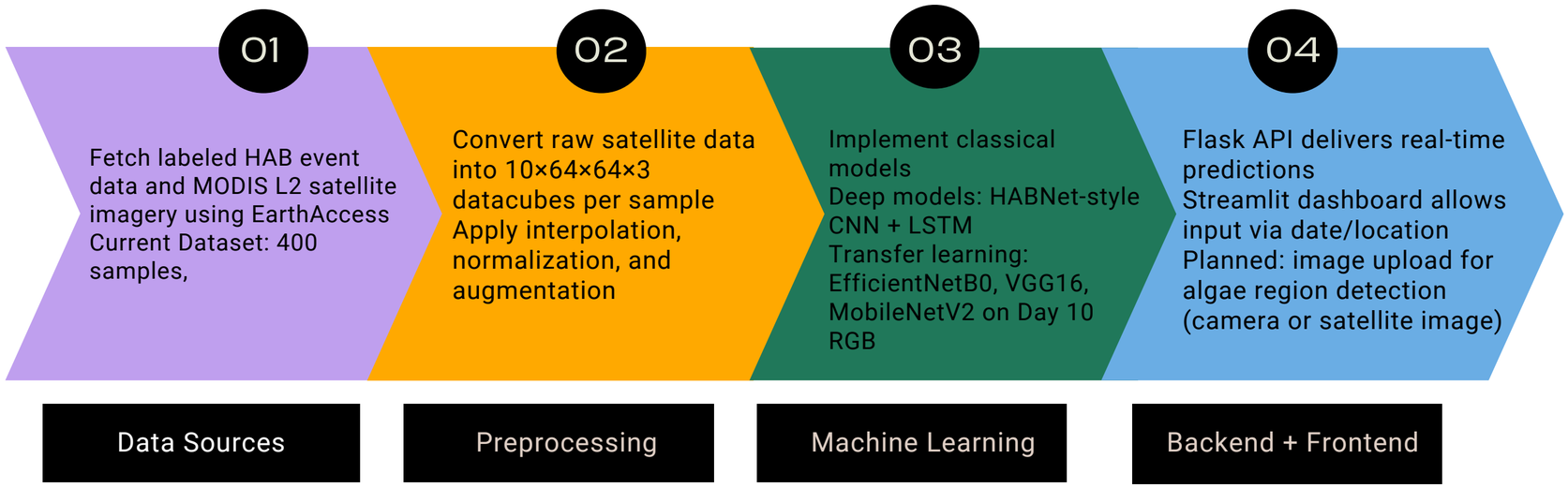
Team: Gradient Descents

Objectives

- Develop a machine learning pipeline to classify HAB events as Toxic or Non-Toxic using multi-day, multi-modality satellite imagery (10-day sequences, 3 modalities)
- Train and evaluate diverse models, including classical ML , deep learning, and transfer learning
- Analyze temporal-spatial patterns and use augmentation to enhance feature learning and interpretability
- Begin integrating real-time capabilities, including prediction interface development and groundwork for algae region detection from uploaded imagery



System Architecture

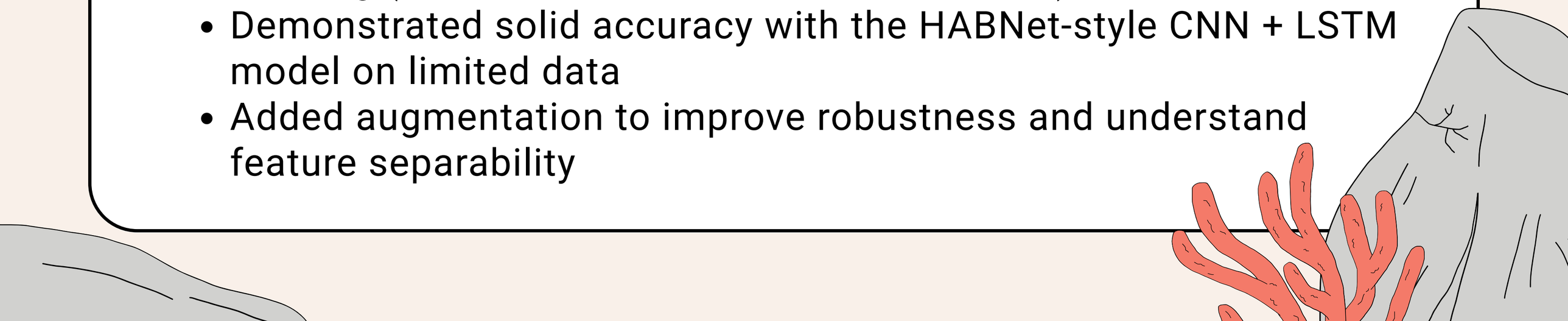


Additions:

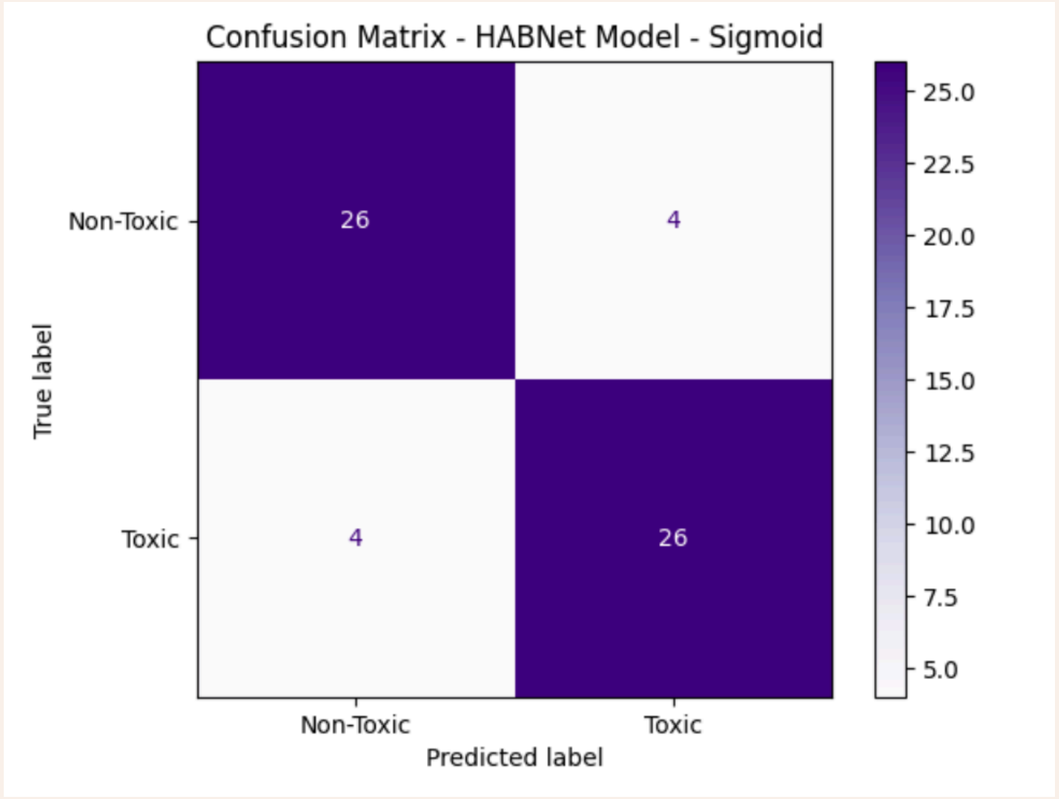
Interactive dashboard connects with Flask API and supports both event-based and manual inputs

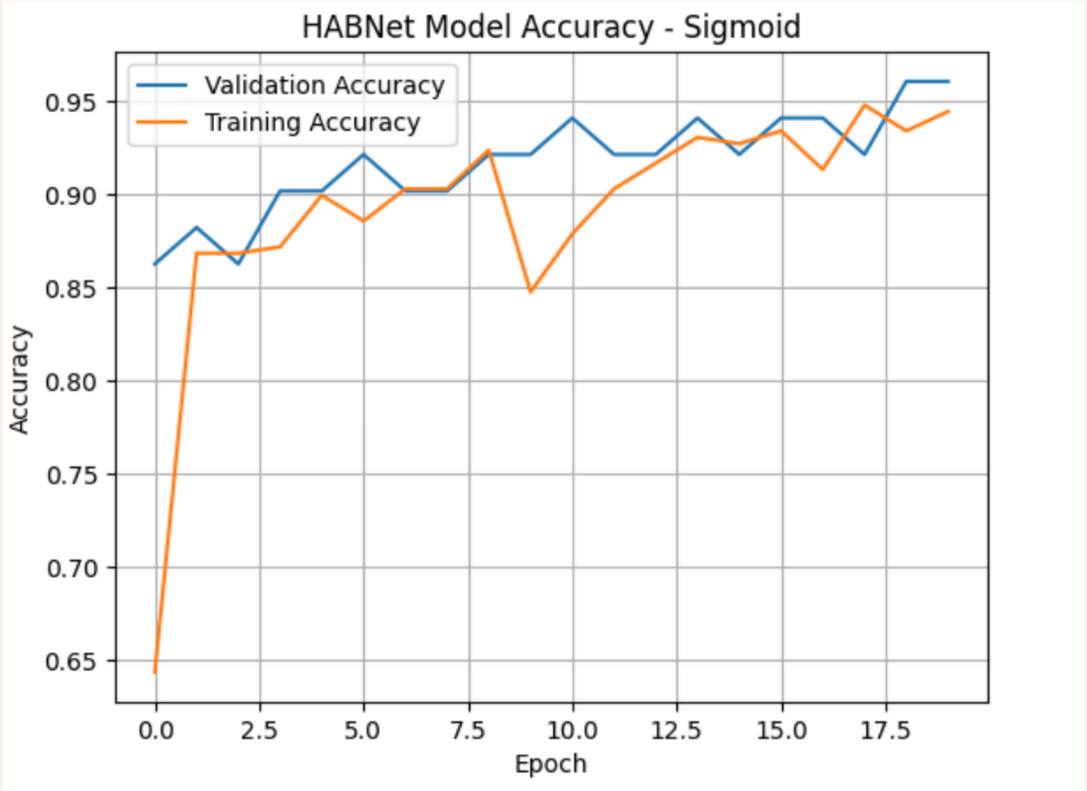
Progress Since Last Presentation

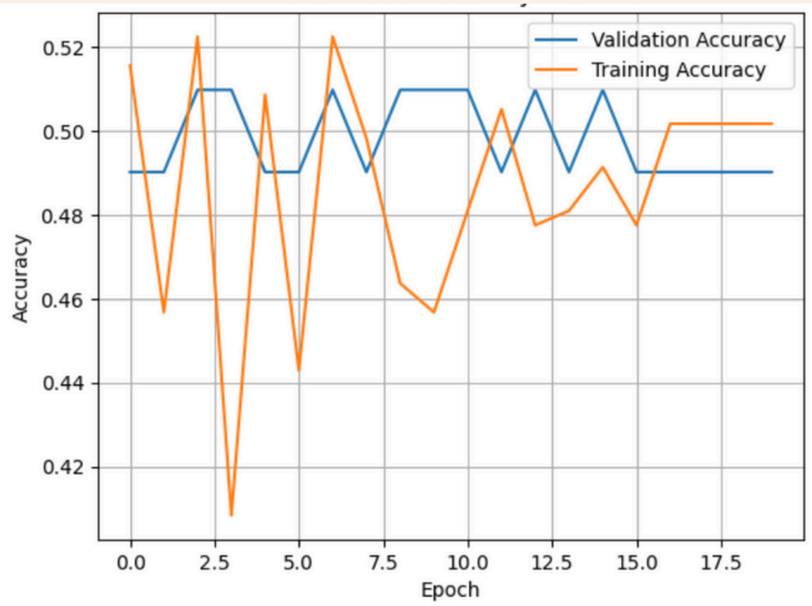
- Developed a full ML pipeline to classify HAB events as Toxic or Non-Toxic using 10-day, 3-modality satellite imagery
- Trained multiple models: baseline ML (Logistic Regression, Ridge, XGBoost), deep learning (CNN + LSTM), and transfer learning (EfficientNetB0, VGG16, MobileNetV2)
- Demonstrated solid accuracy with the HABNet-style CNN + LSTM model on limited data
- Added augmentation to improve robustness and understand feature separability



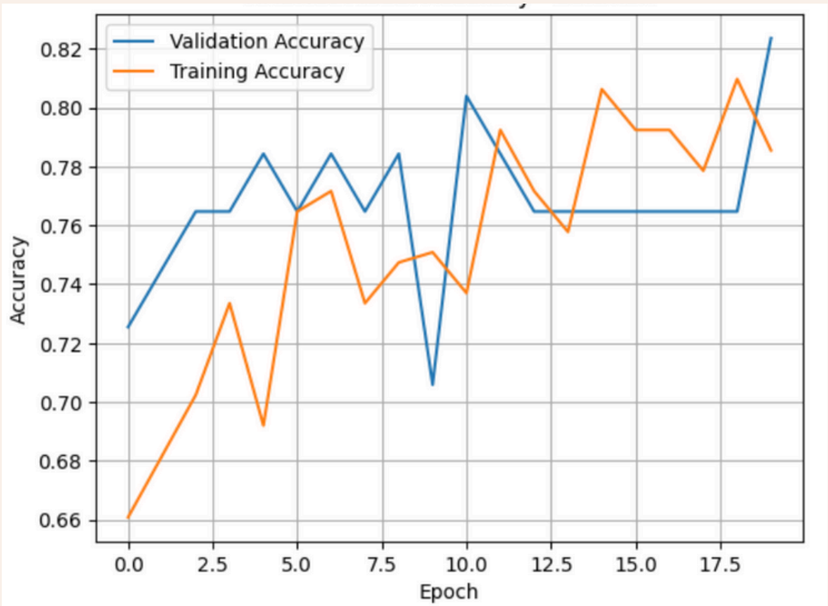
Machine Learning



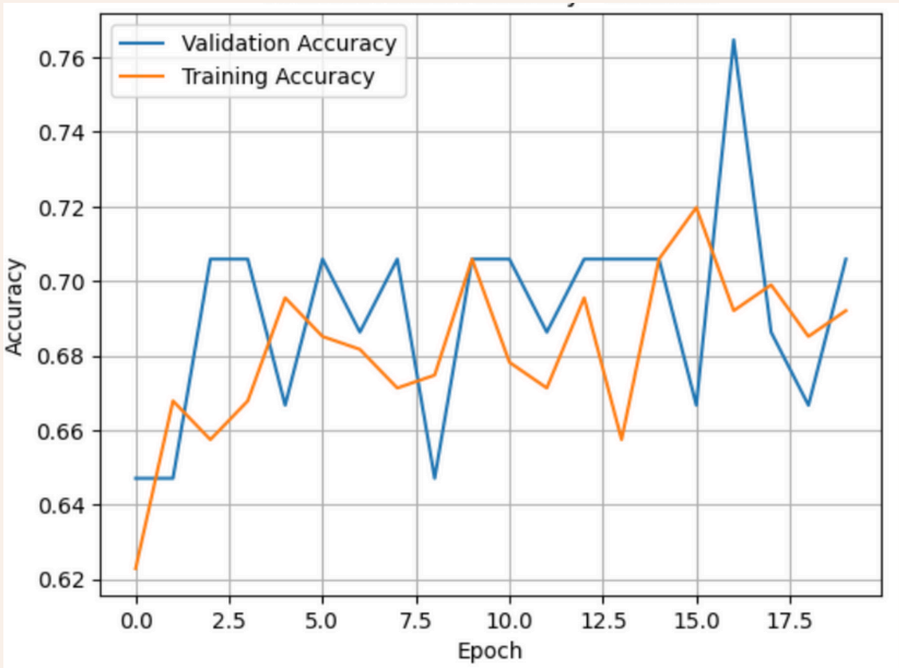




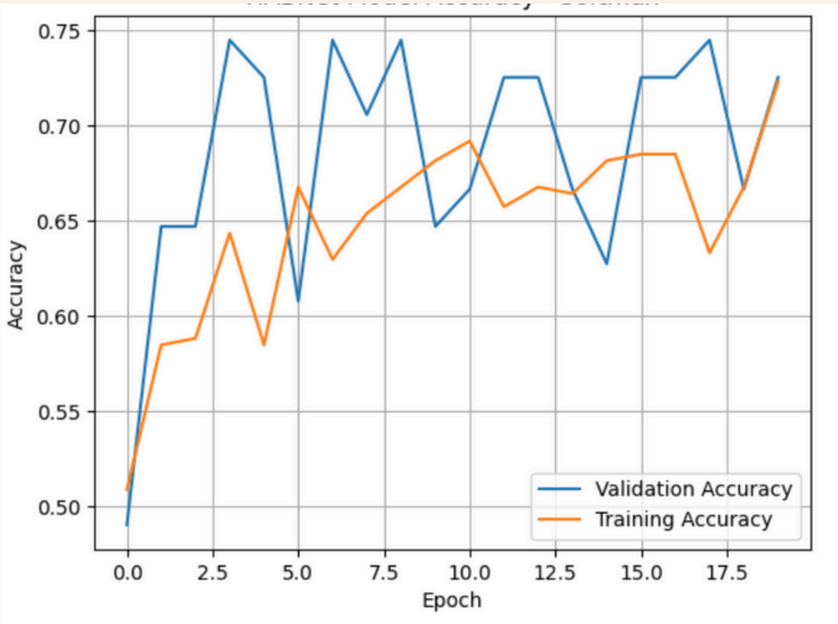
EfficientNetB0



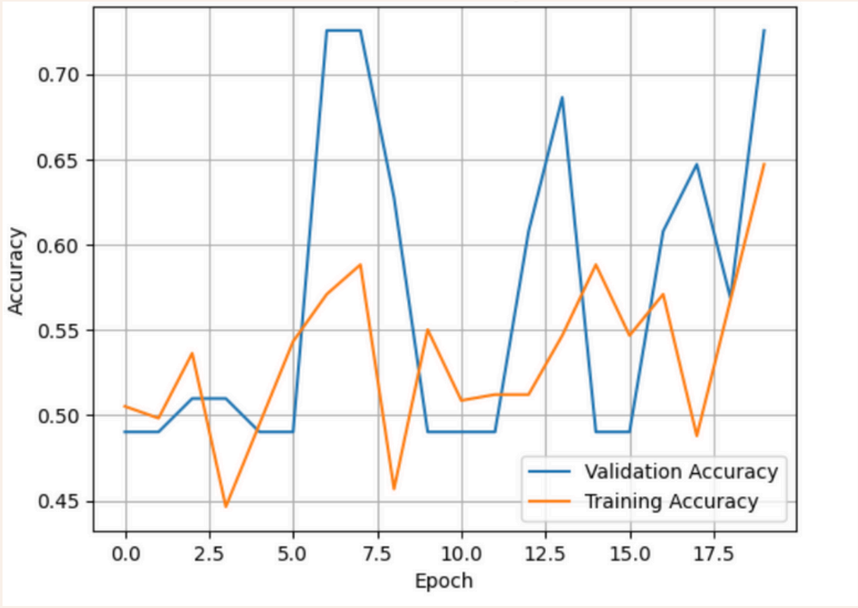
MobileNetV2



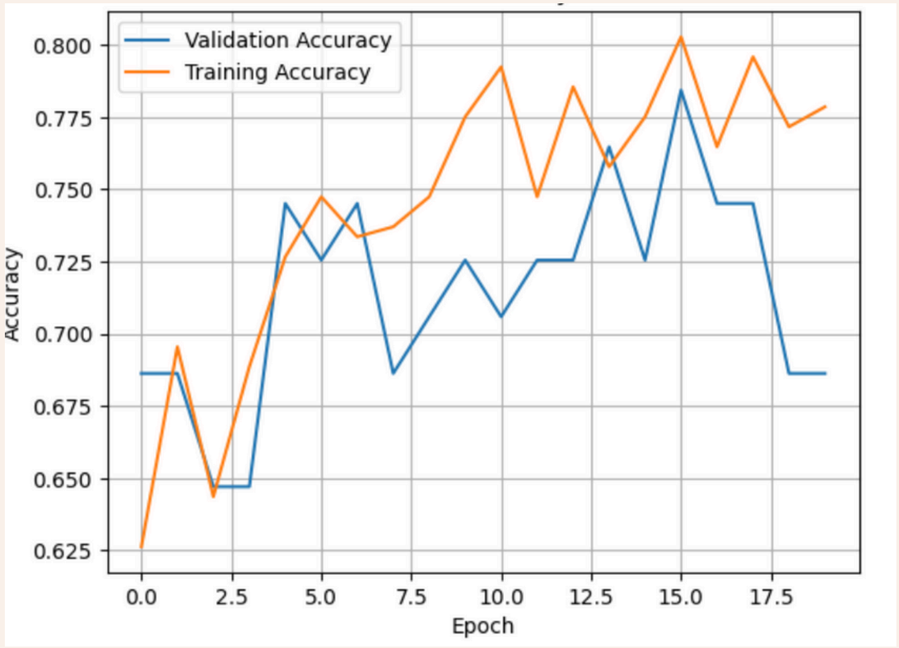
VGG16



VGG19



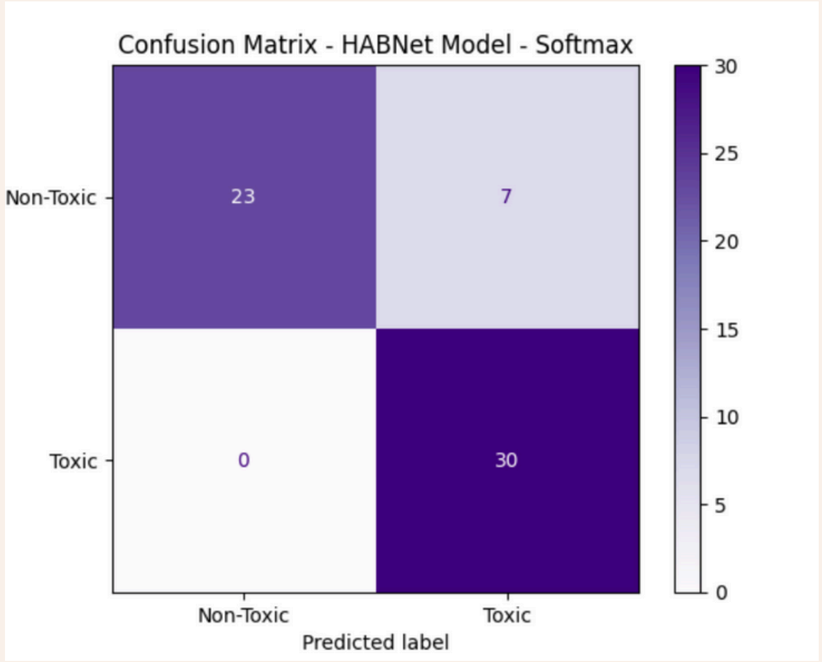
ResNet50

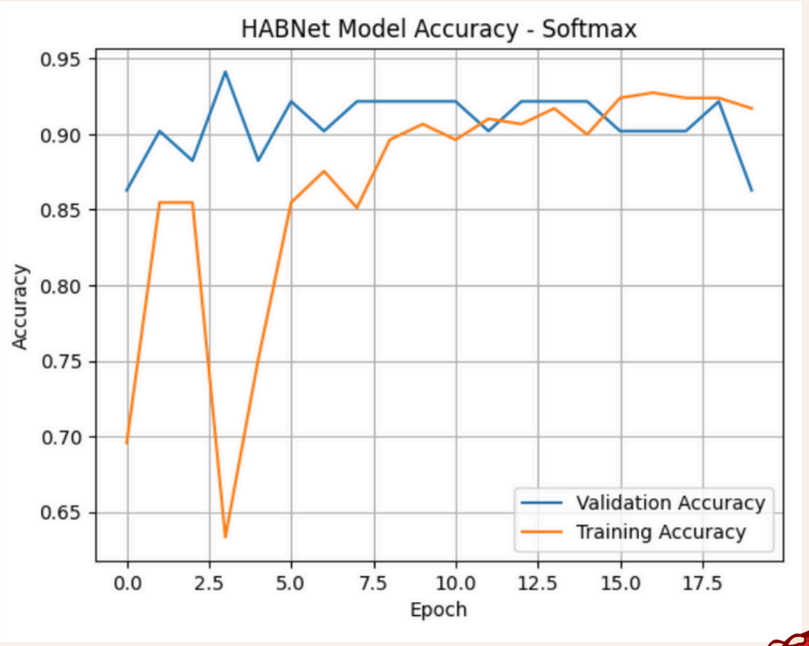


NASNetMobile

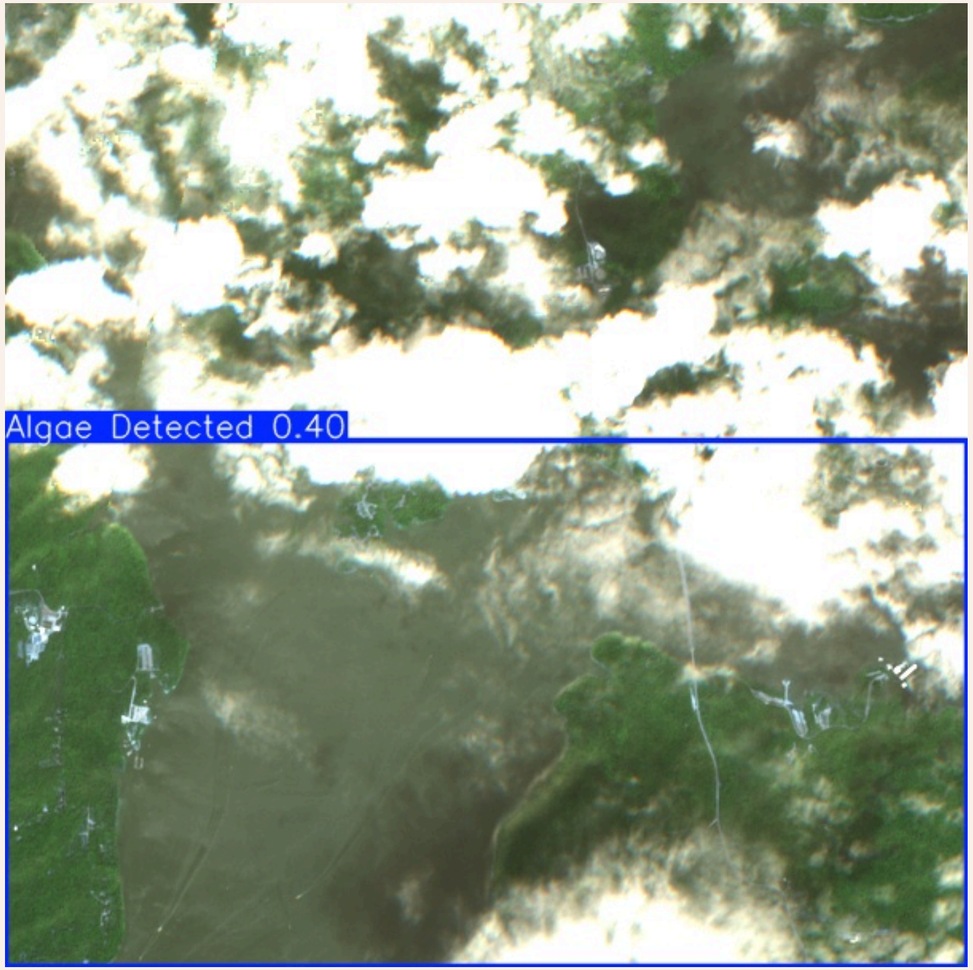
Machine Learning

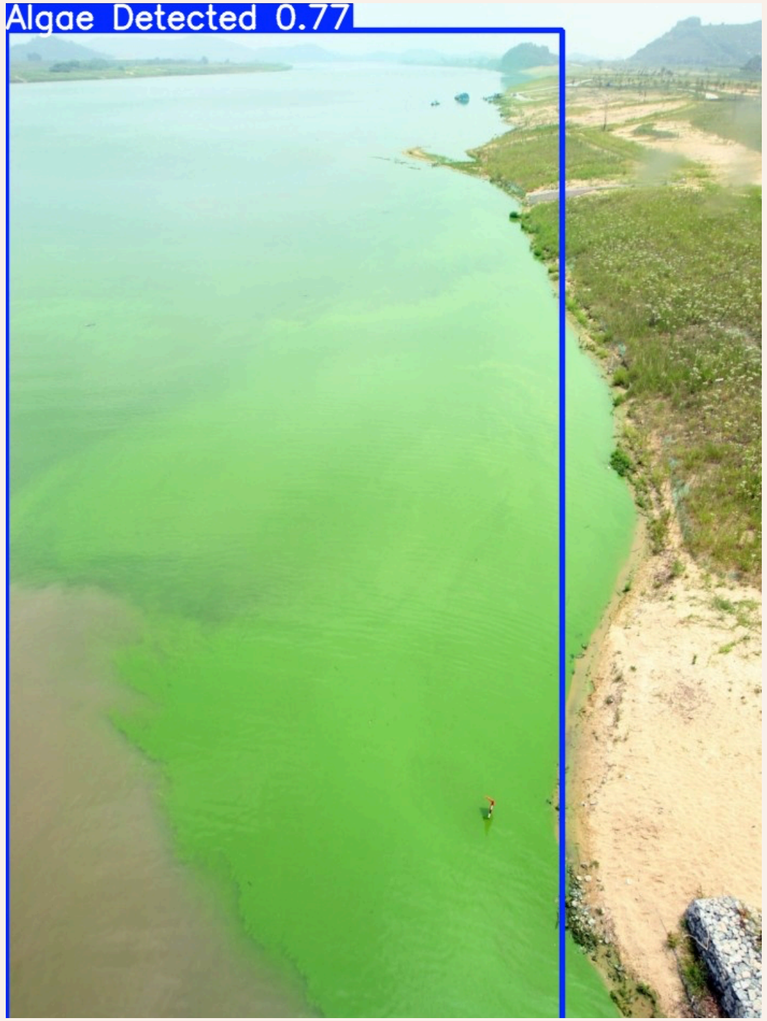
	Model	Accuracy
0	Logistic Regression	0.92
1	Ridge Classifier	0.90
2	XGB Classifier	0.85
3	DT Classifier	0.85
4	RF Classifier	0.88
5	HABnet Sigmoid	0.87
6	HABnet Softmax	0.88
7	EfficientNet	0.50
8	MobileNet	0.77
9	VGGG16	0.75
10	VGG19	0.73
11	ResNet	0.73
12	NasNet	0.70





Added Interactivities





Plans for Coming Week

PIPELINE, BACKEND & FRONTEND

- Optimize the end-to-end workflow to reduce overhead in data extraction and prediction latency
- Deploy the full pipeline on a cloud platform, preferably AWS, for scalable access
- Begin adding image upload functionality to the frontend for detecting algae regions from both camera and satellite images

ML EXPERIMENTS

- Focus on reducing false negatives by improving recall in toxic bloom classification
- Experiment with pre-trained CNNs as feature extractors and combine them with a custom classification neural network
- Compare performance against existing models (e.g., HABNet, XGBoost) for generalization and robustness

LET'S CHECK

Overall Accomplishments

- Developed an end-to-end ML pipeline for classifying HAB events as Toxic or Non-Toxic using 10-day, 3-modality satellite datacubes
- Implemented and evaluated multiple models: baseline classifiers (Logistic Regression, Ridge, XGBoost), deep learning (HABNet-style CNN + LSTM), and transfer learning (EfficientNetB0, VGG16, MobileNetV2)
- Demonstrated strong performance from CNN + LSTM on limited data and high accuracy (>85%) using Day-10 RGB with pretrained CNNs
- Used augmentation for improved model robustness and better class separability
- Laid groundwork for real-time prediction interface, enabling future user interaction and field usage

Review and Tasks Assigned

CHALLENGE

TASKS	ASSIGNEE
Data collection and preprocessing	Daniel
Machine learning and Model Training	Kruthi, Dharmik
Backend	Sagar, Karthika
Frontend	Roshan

Thank You